

## BRIEF COMMUNICATION

# Urine mRNA to identify a novel pseudoexon causing dystrophinopathy

Layal Antoury<sup>1,2</sup>, Ningyan Hu<sup>1,2</sup>, Basil Darras<sup>2,3</sup> & Thurman M. Wheeler<sup>1,2</sup><sup>1</sup>Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts<sup>2</sup>Harvard Medical School, Boston, Massachusetts<sup>3</sup>Department of Neurology, Boston Children's Hospital, Boston, Massachusetts**Correspondence**

Thurman M. Wheeler, Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts. Tel: 617-726-3642; Fax: 617-724-1537; E-mail: twheeler1@mgh.harvard.edu

**Funding Information**

Muscular Dystrophy Association (T.M.W.).

Received: 8 March 2019; Accepted: 20 March 2019

*Annals of Clinical and Translational Neurology* 2019; 6(6): 1106–1112

doi: 10.1002/acn3.777

**Introduction**

Pre-mRNA splicing is the removal of introns followed by the joining together of exons to produce mRNA. The boundaries of exons and introns are demarcated by conserved 5' splice site (5' ss) and 3' splice site (3' ss) intron sequence elements that are recognized by splicing proteins.<sup>1</sup> Approximately 10–15% of annotated pathogenic variants are located in intron splice sites and can cause pathology by inappropriate removal of the adjacent exon, inclusion of intron sequence, or activation of weak pseudo splice sites, leading to a protein product that is dysfunctional, reduced, or absent altogether, even in the setting of a normal coding sequence.<sup>1</sup>

Dystrophinopathies result from pathogenic variants in the X-linked *DMD* (italics) gene and broadly are characterized as two types: (1) Duchenne muscular dystrophy (DMD, no italics), a severe and uniformly fatal disorder, typically in the 20s and (2) Becker muscular dystrophy (BMD), a less severe form with a clinically wide spectrum of phenotypes, ranging from ambulatory to nonambulatory, and a life expectancy from mid-to-late adulthood.<sup>2</sup> DMD is the most common muscular dystrophy of childhood, affecting approximately 1 in 5000 live male births.<sup>3</sup> BMD is about 1/3 as common as DMD.<sup>4</sup> More than 90% of the sequence variants causing DMD are deletions,

**Abstract**

In muscular dystrophies, identification of pathogenic pseudoexons involves sequencing of the target gene cDNA derived from muscle mRNA. Here we use a urine “liquid biopsy,” droplet digital PCR, and sequencing of PCR products to identify a novel cryptic splice site in *DMD* intron 67 that causes dystrophinopathy. Pseudoexon inclusion is 35% in urine cells, 34% in urine extracellular RNA (exRNA), and 54% in muscle biopsy tissue, but absent in serum exRNA. Our results suggest that cryptic splice site use varies depending on the RNA source, and that urine RNA has the potential to substitute for muscle biopsies to identify *DMD* pseudoexons.

uplications, or premature termination codons that interrupt the open reading frame (ORF) and result in the absence of functional dystrophin protein.<sup>5,6</sup> By contrast, BMD most often is due to sequence variants in the *DMD* gene that maintain the ORF, resulting in an internally truncated partially functional dystrophin protein that has a preserved carboxy-terminus.<sup>5,7</sup>

Diagnosis of DMD and BMD by analysis of genomic DNA from peripheral blood lymphocytes identifies >90% of *DMD* sequence variants.<sup>6</sup> In recent years, it largely has replaced the need for muscle biopsy to establish the diagnosis of dystrophinopathy. However, approximately 7% of sequence variants that cause dystrophinopathy are located in noncoding regions,<sup>8</sup> some of which induce a cryptic splice site that leads to aberrant inclusion of intron sequence in the mRNA known as a pseudoexon.<sup>6,9,10</sup> Identification of pathogenic *DMD* pseudoexons generally requires sequencing of cDNA derived from mRNA of muscle tissue or muscle cells.<sup>6</sup>

Extracellular vesicles (EVs) are membrane-encased particles that are released from cells and carry nucleic acids, proteins, and lipids.<sup>11,12</sup> EVs may play an important role in cell-to-cell communication in normal and disease states, and have been examined as a novel source of biomarkers in blood, urine, and CSF.<sup>13–15</sup> The RNA

within EVs includes mRNAs, micro RNAs (miRNAs), and noncoding RNAs, collectively termed extracellular RNA (exRNA). Recently, we found that exRNA in human urine contains splice variant biomarkers of myotonic dystrophy and patient-specific *DMD* deletion transcripts that can demonstrate activity of the splice shifting drug eteplirsen.<sup>16</sup> In this study, we examined whether exRNA also could be used to determine whether an intron base substitution in the *DMD* gene disrupts splicing efficiency as a mechanism of dystrophin-deficient muscular dystrophy.

## Materials and Methods

### Human subjects

The Massachusetts General Hospital/Partners Health Service Institutional Review Board (IRB) approved this study. We recruited study participants from the Massachusetts General Hospital Pediatric Neuromuscular Clinic. Prior to participation in the study, both subjects provided informed consent. Due to severe autism in the individual with Becker muscular dystrophy (BMD), his mother/legal guardian provided informed consent for his participation, according to IRB protocol.

### Microscopy and image capture

We reviewed the diagnostic biceps muscle biopsy slides prepared by Boston Children's Hospital Department of Pathology as part of routine clinical care, which included frozen muscle sections stained with hematoxylin and eosin (H&E) or labeled using antidystrophin antibodies targeting the N-terminus (DYS3), rod domain (DYS1), or carboxy-terminus (DYS2). To capture images, we used an AxioImager microscope (Zeiss), 10× and 20× objectives, a MicroPublisher 3.3 RTV color CCD camera (Q-Imaging), and Volocity image acquisition software (Perkin Elmer).

### Biofluid collection and processing

We collected voided urine samples in a standard urine specimen cup, and blood samples in two red top serum separator tubes. Urine volumes were 58 mL from the BMD individual and 90 mL from the unaffected (UA) individual. Serum volumes were 5.9 mL (BMD) and 7 mL (UA). The BMD and UA blood and urine samples were collected on the same day and processed in parallel to remove cells, as previously described.<sup>16</sup>

### Isolation of exRNA and total RNA

We ultracentrifuged urine and serum samples at 100,000g and extracted exRNA from the ribonucleoprotein pellet

using Trizol (Invitrogen) according to manufacturer recommendations, as previously described.<sup>16</sup> We also used Trizol to isolate total RNA from urine cells and muscle biopsy tissue.

### Splicing analysis by RT-PCR and droplet digital PCR (ddPCR)

We used random primers and Superscript III to generate cDNA, and examined *DMD* intron 67 splicing using RT-PCR, Sanger sequencing of PCR products, and ddPCR using gene-specific primers, as previously described.<sup>16</sup>

RT-PCR primers (Invitrogen; normal splice product size 267 bp):

Left primer: 5'-ACTGGCATCATTTCCTGTG-3'  
Right primer: 5'-GGTTCAGTCTCATCCAGTC-3'

ddPCR primer probe sets (Biosearch Technologies):

DMD-67int-68 (amplicon size 102 bp):

Left primer: 5'-GTAACATTGAGCCAAGTGTC-3'  
Right primer: 5'-GGTTCAGTCTCATCCAGTC-3'  
Probe: 5'-FAM-CCAATTTgtaagGCTAATAATAAGC  
C-BHQ1-3'

DMD-67-68 (amplicon size 97 bp):

Left primer: 5'-GTAACATTGAGCCAAGTGTC-3'  
Right primer: 5'-GGTTCAGTCTCATCCAGTC-3'  
Probe: 5'-FAM-TTCCAATTTGCTAATAATAAGCCA-B  
HQ1-3'

## Results

### Dystrophinopathy associated with an intron base substitution in the *DMD* gene

A 5-year old boy with severe autism and serum creatine kinase values of 20,080 and 11,710 U/L one week apart was diagnosed with DMD after a quadriceps muscle biopsy reportedly showed absence of dystrophin protein by both immunolabeling of muscle tissue sections and Western blotting. At age six, using DNA isolated from blood, a comprehensive mutation analysis of the *DMD* gene amplified in 85 segments covering 20.5 kb, including all of the coding regions and associated splice junctions, demonstrated a normal coding sequence and a single base substitution (c9807+ 6 T>G) in intron 67 (Fig. 1A). At age 12, he stopped walking shortly after a fall without serious injury, suggesting either an unexpectedly sudden and rapid progression of weakness or that some component of autism may have played a role. At age 19, his arm strength was atypical of DMD and

included the ability to lift both of his arms above his head, raising questions about the diagnosis. Due to the unavailability of his original quadriceps muscle biopsy for repeat testing, he underwent a biceps muscle biopsy, his second diagnostic muscle biopsy, which confirmed a severe dystrophic myopathy. However, unlike his quadriceps muscle biopsy at age 5, immunohistochemistry using antibodies targeting the amino-terminal, rod domain, and the carboxy-terminal portions of the protein all demonstrated that dystrophin protein was present, although reduced, in his biceps (Fig. 1B). Expression of dystrophin protein in this patient is inconsistent with the diagnosis of DMD, but would be compatible with BMD. The presence of dystrophin protein expression in his second muscle biopsy also prompted discussions of whether the base substitution in intron 67 is pathogenic, or whether a variant in a different gene should be considered.

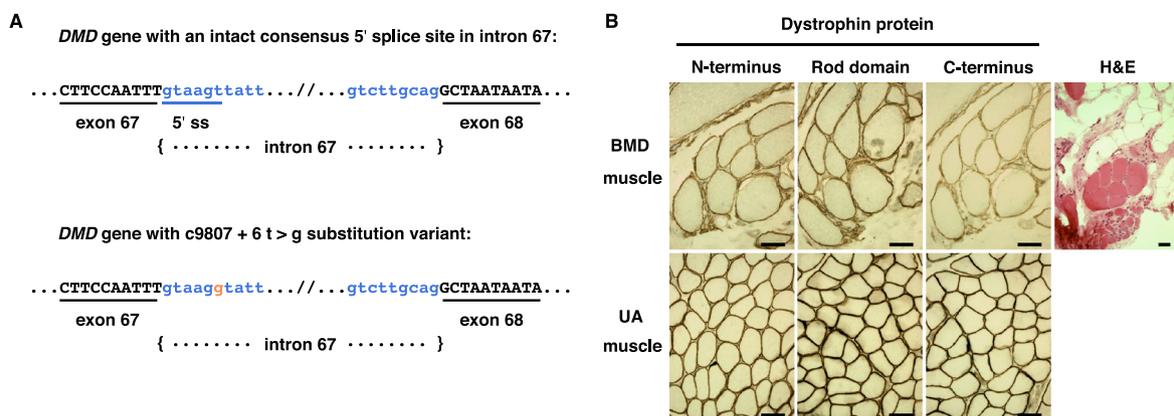
### Identification of pseudoexon inclusion using urine RNA

The normal *DMD* coding sequence in this individual would be expected to produce a normal dystrophin protein. To examine the possibility that the base substitution in intron 67 may cause dystrophinopathy by abnormal RNA splicing, we collected urine and serum samples from this individual at age 20 to examine *DMD* mRNA expression for a potential pseudoexon. Biofluid specimens from his UA mother, who is a noncarrier for the base substitution, and commercially available skeletal muscle total RNA served as controls.

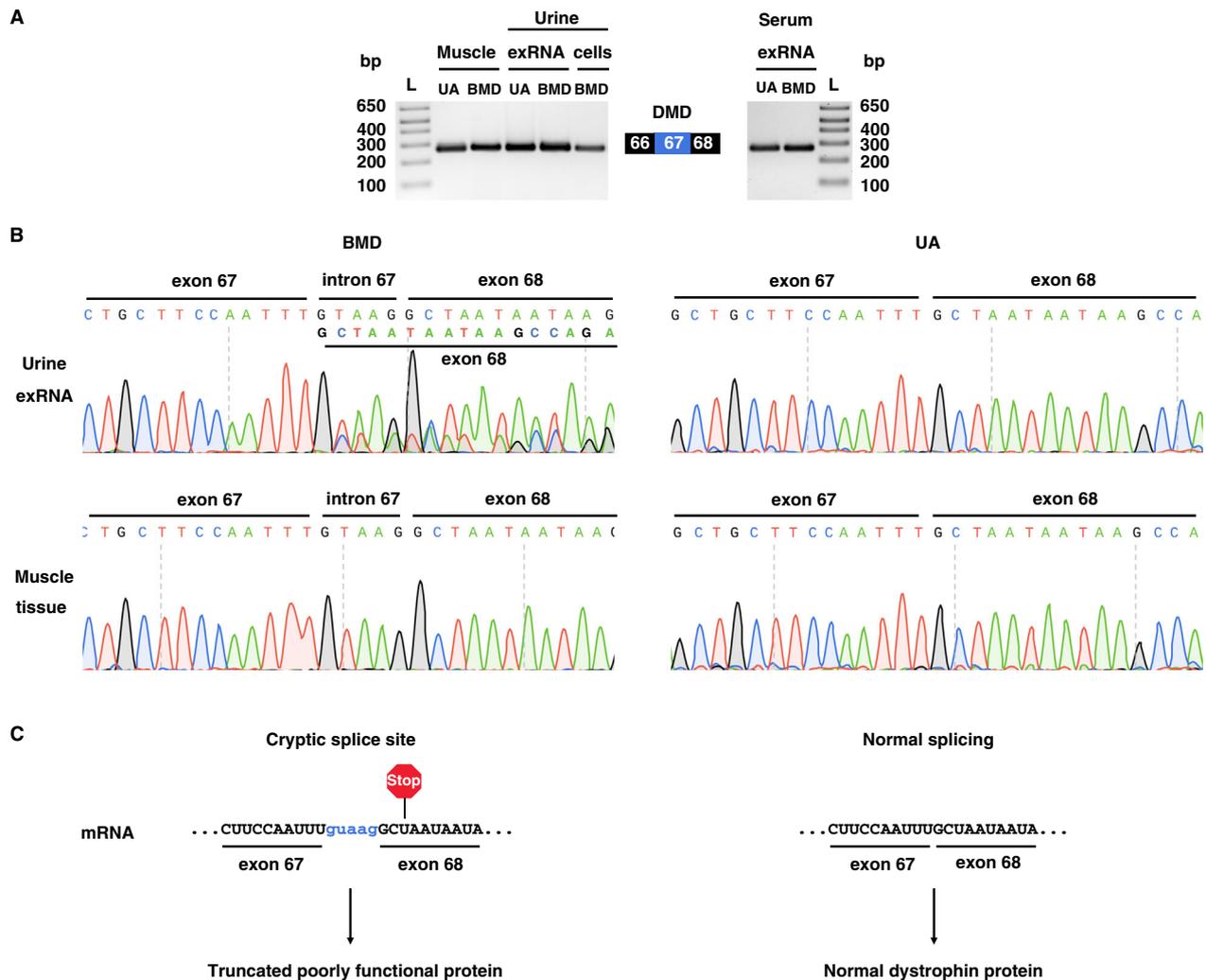
Using primers targeting *DMD* exons 66 and 68, RT-PCR analysis revealed a single splice product in urine exRNA, urine cells, and serum exRNA specimens from the BMD subject that appeared similar to the splice product in UA control biofluid and muscle tissue (Fig. 2A). Sequencing of PCR products from UA urine exRNA and muscle tissue confirmed the expected normal transition from exon 67 to 68 (Fig. 2B). However, sequencing of the urine exRNA PCR product from the BMD subject revealed a splice variant that includes the first five nucleotides of intron 67 (Fig. 2C). Examination of the chromatogram also indicated a second PCR product containing the normal transition from exon 67 to 68 that is identical to the product in UA urine exRNA and muscle tissue (Fig. 2C). To confirm the pathogenic effects of this cryptic splice site, we next examined total RNA in the remaining biceps muscle biopsy tissue from this individual, and found a single splice product that also contains the intron 67 inclusion element, identical to the urine exRNA sample from this individual. Inclusion of the five nucleotide intron sequence shifts the ORF, resulting in a premature termination codon in exon 68 (Fig. 2C).

### Quantitation of splicing efficiency in mRNA

Although sequencing of the PCR product from his biceps muscle biopsy showed a pseudoexon that causes a premature termination codon, the dystrophin protein evident in his biceps muscle biopsy suggests that at least some of the *DMD* transcripts are spliced normally, which would be expected to produce a normal



**Figure 1.** A novel point mutation in the 5' splice site of intron 67 of the *DMD* gene causes dystrophinopathy. (A) Sequence of a normal *DMD* gene with an intact consensus 5' splice site (5' ss) in intron 67 (upper) and the location of the *DMD* intron 67 t-to-g substitution variant, shown in orange, that disrupts the 5' ss (lower). (B) Immunohistochemical analysis of biceps muscle biopsy cryosections from the individual with BMD (upper) using antibodies specific for the N-terminus, rod domain, and C-terminus portions of the dystrophin protein. A normal human muscle biopsy served as a control (lower). Hematoxylin and eosin (H&E) shows dystrophic muscle characterized by extensive loss of muscle fibers and replacement with fat and connective tissue. Bars = 50  $\mu$ m.

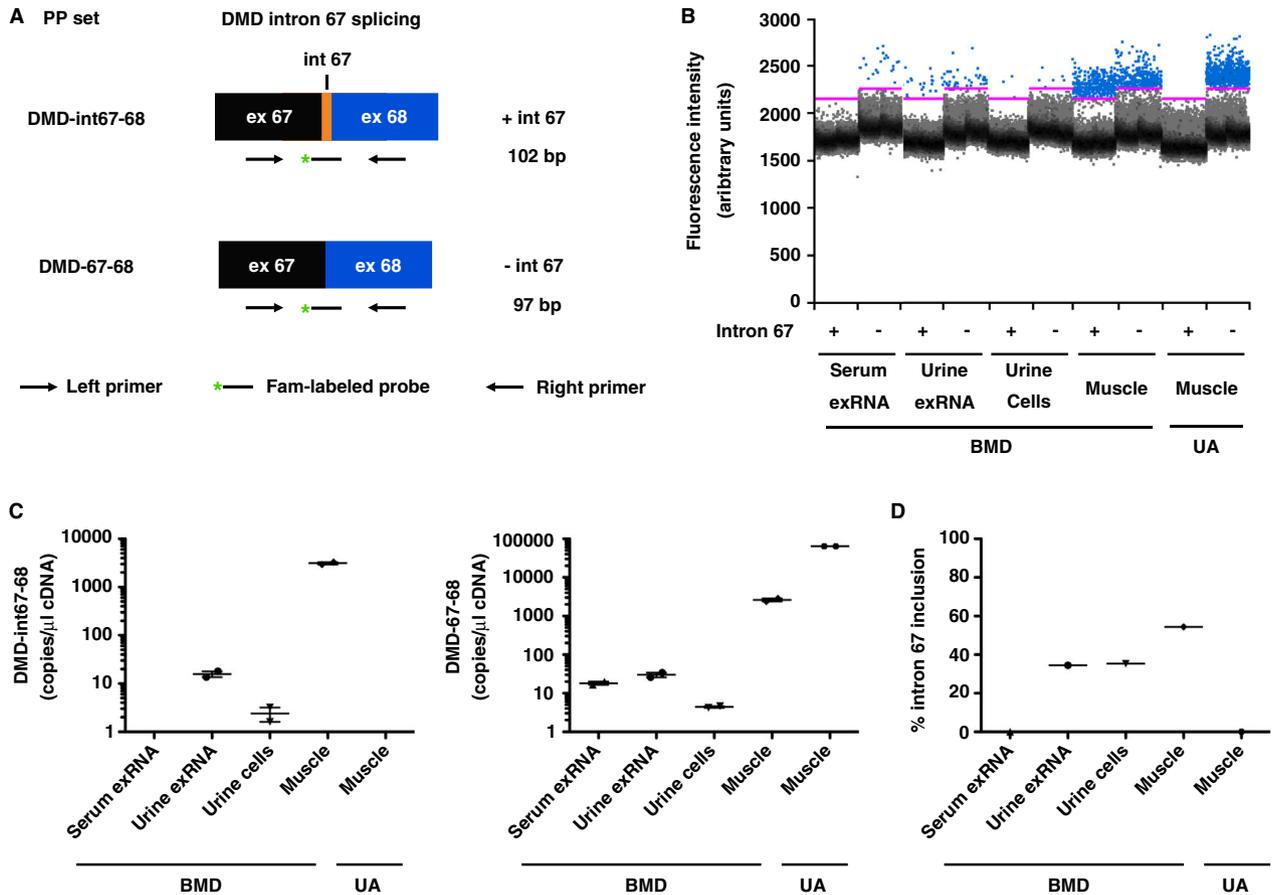


**Figure 2.** Urine exRNA identifies a *DMD* cryptic splice site and frame-shifting pseudoexon. (A) Examination of urine exRNA, urine cell total RNA, serum exRNA, and commercially available total RNA from UA muscle tissue using RT-PCR and primers targeting *DMD* exons 66 and 68. (B) Sequencing of the PCR products from UA urine exRNA and muscle (right upper and lower) showed normal splicing of exons 67 and 68. By contrast, splice products in the urine exRNA and biceps muscle biopsy (left upper and lower) from the BMD patient both show inclusion of the first five nucleotides of intron 77. In addition, examination of the urine exRNA chromatogram also reveals a second splice product that shows the normal transition from exon 67 to exon 68, identical to UA urine exRNA and muscle. (C) Diagram of the cryptic splice site in intron 77 that shifts the reading frame to create a premature termination codon in exon 68, explaining the dystrophinopathy in this individual.

dystrophin protein. Droplet digital PCR (ddPCR) enables precise quantification of *DMD* splice products in human urine exRNA, human urine and muscle cells, and mouse muscle tissue.<sup>16–18</sup> Using ddPCR and primer probe sets that are specific for either the normal splice event or pseudoexon inclusion, we found that use of the cryptic splice site was 34% of transcripts in urine exRNA, 35% in urine cells, 54% in biceps muscle biopsy tissue, and absent in serum exRNA (all transcripts in serum exRNA were spliced normally) (Fig. 3).

## Discussion

Recently we reported that exRNA in human urine contains splice variant biomarkers of disease activity in myotonic dystrophy, and patient-specific *DMD* deletion transcripts in Duchenne muscular dystrophy.<sup>16</sup> Here we expand the use of urine exRNA to define a novel *DMD* cryptic splice site and a frame-shifting pseudoexon in an individual with dystrophin-deficient muscular dystrophy despite a normal *DMD* coding sequence. Until now, mRNA isolated from biopsy material has been required to



**Figure 3.** Quantification of intron 67 inclusion by droplet digital PCR (ddPCR). We used ddPCR to examine *DMD* intron 67 inclusion in serum exRNA, urine exRNA, urine cells, and biceps muscle biopsy tissue samples. Total RNA from UA muscle tissue served as a control. (A) Assay design using separate primer probe (PP) sets to identify intron 67 inclusion (DMD-int67-68) and intron 67 exclusion (DMD-67-68). The left primers targeting exon 67 and the right primers targeting exon 68 are identical in each PP set. Identification of intron 67 (orange)-containing transcripts uses a fluorescent FAM-labeled probe that targets the splice junction for exon 67-intron 67-exon 68, while identification of normally spliced transcripts uses a FAM-labeled probe that targets the splice junction for exons 67-68. The DMD-int67-68 PP set amplifies only those transcripts that contain intron 67, and has an amplicon size of 102 base pairs (bp). The DMD-67-68 PP set amplifies only those transcripts that are spliced normally, lack intron 67, and have an amplicon size of 97 bp. (B) Representative droplet populations (blue) in BMD serum exRNA, urine exRNA, urine cells, and muscle, and UA muscle samples using each assay (DMD-int67-68 and DMD-67-68). (C) Quantification of *DMD* splice products that include intron 67 (left) and exclude intron 67 (right) as copies per microliter cDNA. Individual data points represent the mean of duplicate assays for each sample. Note that values for intron 67 inclusion in serum exRNA and UA muscle tissue RNA are zero. Error bars indicate mean  $\pm$  s.e.m. (D) Calculation of *DMD* intron 67 inclusion using data in (C). Note that values for serum exRNA and UA muscle are zero.

identify the inclusion of intron sequence. Our results suggest that examination of urine RNA has the potential to replace the need for muscle tissue biopsy to identify pathogenic *DMD* pseudoexons, and may extend to include cryptic splice sites and intron inclusion elements causing other genetic disorders that also require tissue biopsy for diagnosis. Urine samples are noninvasive, painless, easy and inexpensive to collect, and carry no risk of general anesthesia that typically is required for tissue biopsies of pediatric patients.

*DMD* intron 67 contains the highly conserved 5' splice site consensus sequence gtaagt(t...) that is

complementary to the sequence within the U1 small nuclear RNA (snRNA) that initiates the splicing reaction.<sup>1</sup> The intron base substitution in this individual alters the 5' splice site sequence to gtaagg(t...). Although 5' splice site variants often result in removal of the adjacent exon from the mRNA,<sup>1</sup> we found no evidence of exon 67 skipping in any of the mRNA sources from this individual. We propose that the t > g substitution weakens pre-mRNA splicing efficiency by simultaneously disrupting complementarity to the U1 snRNA 5' terminus and creating a de novo canonical 5' gt splice site. The result is normal splicing of some

transcripts, and inclusion of a five nucleotide frame-shifting intron sequence in other transcripts, which combine to reduce the overall dystrophin protein expression. Identification of pathogenic variants that disrupt splicing efficiency may be important to confirm a clinical diagnosis for an individual, improve the understanding of splicing regulation more generally, and aid in the development of new therapies that modulate splicing. A single base substitution may be particularly amenable to therapeutic genome editing approaches that are currently under development.<sup>19,20</sup>

We examined three “liquid biopsy” sources of *DMD* transcripts: urine exRNA, urine cell total RNA, and serum exRNA. The differential pseudoexon inclusion among biofluid samples and muscle biopsy tissue that we observed suggests that the impact of a base substitution on splicing efficiency depends on the source of RNA. The absence of intron inclusion in serum exRNA suggests that urine may be preferable as an exRNA source for identification of *DMD* mis-splicing and pathogenic pseudoexons, and that the primary cell or tissue source that contribute *DMD* transcripts to the serum exRNA pool is unlikely to be muscle. The reported absence of dystrophin protein expression in the first biopsy of his quadriceps muscle suggests the possibility that pseudoexon inclusion in the quadriceps muscle may have been close to 100%, rather than the 54% evident in the biceps muscle biopsy that expresses reduced dystrophin protein, and that splicing efficiency may correlate with disease severity of individual muscles.

## Acknowledgments

We thank the patient and his mother who donated blood and urine samples for this study, the MGH DNA core for sequencing support, and H. Lidov (Boston Children’s Hospital) for providing the biceps muscle biopsy tissue for analysis, and X. O. Breakefield (Massachusetts General Hospital) for critical review of the manuscript.

## Conflict of Interest

A patent application has been filed on the use of urine exRNA as markers of muscular dystrophies.

## Author Contributions

L.A. and T.M.W. obtained informed consent and collected samples. L.A., N.H., and T.M.W. processed samples and performed experiments. T.M.W. designed the study and wrote the paper. L.A., N.H., B.D., and T.M.W. analyzed the data, discussed the results, and commented on the manuscript.

## References

- Havens MA, Duelli DM, Hastings ML. Targeting RNA splicing for disease therapy. *Wiley Interdiscip Rev RNA* 2013;4:247–266.
- Flanigan KM. Duchenne and Becker muscular dystrophies. *Neurol Clin* 2014;32:671–688, viii.
- Mendell JR, Shilling C, Leslie ND, et al. Evidence-based path to newborn screening for Duchenne muscular dystrophy. *Ann Neurol* 2012;71:304–313.
- Bushby KM, Thambiyah M, Gardner-Medwin D. Prevalence and incidence of Becker muscular dystrophy. *Lancet* 1991;337:1022–1024.
- Aartsma-Rus A, Van Deutekom JC, Fokkema IF, et al. Entries in the Leiden Duchenne muscular dystrophy mutation database: an overview of mutation types and paradoxical cases that confirm the reading-frame rule. *Muscle Nerve* 2006;34:135–144.
- Gurvich OL, Tuohy TM, Howard MT, et al. *DMD* pseudoexon mutations: splicing efficiency, phenotype, and potential therapy. *Ann Neurol* 2008;63:81–89.
- Monaco AP, Bertelson CJ, Liechti-Gallati S, et al. An explanation for the phenotypic differences between patients bearing partial deletions of the *DMD* locus. *Genomics* 1988;2:90–95.
- Dent KM, Dunn DM, von Niederhausern AC, et al. Improved molecular diagnosis of dystrophinopathies in an unselected clinical cohort. *Am J Med Genet A* 2005;134:295–298.
- Roest PA, Bout M, van der Tuijn AC, et al. Splicing mutations in *DMD/BMD* detected by RT-PCR/PTT: detection of a 19AA insertion in the cysteine rich domain of dystrophin compatible with *BMD*. *J Med Genet* 1996;33:935–939.
- Tuffery-Giraud S, Saquet C, Chambert S, Claustres M. Pseudoexon activation in the *DMD* gene as a novel mechanism for Becker muscular dystrophy. *Hum Mutat* 2003;21:608–614.
- Tkach M, Thery C. Communication by extracellular vesicles: where we are and where we need to go. *Cell* 2016;164:1226–1232.
- Erdbrugger U, Le TH. Extracellular vesicles in renal diseases: more than novel biomarkers? *J Am Soc Nephrol* 2016;27:12–26.
- Skog J, Wurdinger T, van Rijn S, et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat Cell Biol* 2008;10:1470–1476.
- Nilsson J, Skog J, Nordstrand A, et al. Prostate cancer-derived urine exosomes: a novel approach to biomarkers for prostate cancer. *Br J Cancer* 2009;100:1603–1607.
- Chen WW, Balaj L, Liau LM, et al. BEAMing and droplet digital PCR analysis of mutant *IDH1* mRNA in glioma

- patient serum and cerebrospinal fluid extracellular vesicles. *Mol Ther Nucleic Acids* 2013;2:e109.
16. Antoury L, Hu N, Balaj L, et al. Analysis of extracellular mRNA in human urine reveals splice variant biomarkers of muscular dystrophies. *Nat Commun* 2018;9:3906.
  17. Verheul RC, van Deutekom JC, Datson NA. Digital droplet PCR for the absolute quantification of exon skipping induced by antisense oligonucleotides in (pre-) clinical development for Duchenne muscular dystrophy. *PLoS ONE* 2016;11:e0162467.
  18. Hiller M, Spitali P, Datson N, Aartsma-Rus A. Exon 51 skipping quantification by digital droplet PCR in del52hDMD/mdx mice. *Methods Mol Biol* 2018;1828:249–262.
  19. Bertoni C, Rustagi A, Rando TA. Enhanced gene repair mediated by methyl-CpG-modified single-stranded oligonucleotides. *Nucleic Acids Res* 2009;37:7468–7482.
  20. Nelson CE, Robinson-Hamm JN, Gersbach CA. Genome engineering: a new approach to gene therapy for neuromuscular disorders. *Nat Rev Neurol* 2017;13:647–661.